Genomic Medicine, Biomarkers, and Health Sciences (2012) 4, 3–6

Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.e-gmbhs.com

INVITED SHORT COMMUNICATION

Combining MALDI-TOF and molecular imaging with principal component analysis for biomarker discovery and clinical diagnosis of cancer

Yi-Tzu Cho [a], Yi-Yan Chiang [b], Jentaie Shiea [b,*], Ming-Feng Hou [c]

[a] *Department of Cosmetic Applications and Management, Yuh-Ing Junior College of Health Care & Management, Kaohsiung, Taiwan*
[b] *Department of Chemistry, National Sun Yat-Sen University, Kaohsiung, Taiwan*
[c] *Cancer Research Center, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan*

Received 24 April 2012; accepted 24 April 2012
Available online 12 May 2012

KEYWORDS
matrix-assisted laser desorption/ ionization mass spectrometry;
molecular imaging;
principle component analysis;
tissue;
diagnosis

Abstract Molecular imaging using matrix-assisted laser desorption ionization/time-of-flight mass spectrometry (MALDI-TOF MS) is effective for determining the distribution of molecules of interest in specific tissues. It can determine the direct correlation between metabolite, lipid, and protein expression and histology. Principle component analysis (PCA) can reduce the dimensions of a data set while still retaining the information present in the original data set. Using PCA to process MALDI data, samples with different statuses and ion patterns on their MALDI mass spectra can be classified, grouped, and evaluated on the same score plot. The use of MALDI-TOF in combination with PCA to compare the lipid, peptide, and protein profiles of different biological specimens can then be used to diagnose disease. Because ions with significant differences between sampling regions in a tissue can be indicated using PCA, the imaging of these "interesting peaks" can be visualized by plotting the ion intensity across the tissue section.

Introduction

Mass spectrometry has advanced to be a superior tool for protein identification in proteomic strategies. It often combines with multidimensional prefractionation techniques, due to the high dynamic range of protein concentrations in the complex matrix, protolytic digestion and database searching for discovering disease biomarkers. However, very few biomarkers discovered by these strategies have been successfully introduced into clinical practice. Each of these laborious sample pretreatment steps suffers from lack of reproducibility and consistency.

* Corresponding author. Department of Chemistry, National Sun Yat-Sen University, No. 70, Lienhai Rd., Kaohsiung 80424, Taiwan.
E-mail address: jetea@mail.nsysu.edu.tw (J. Shiea).

2211-4254/$36
doi:10.1016/j.gmbhs.2012.04.022

Since mass spectrometry owns the innate features of accuracy, sensitivity, high speed, and automation; this makes it be a potential tool for performing high-throughput diagnostic analyses. Herein, we describe matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry combined with molecular imaging and multivariate analysis as an efficient approach to directly characterize potential biomarkers on tissue samples. Principal component analysis (PCA) was used to process mass spectra of lipid, peptide, protein, or imaging data from bio-specimens to rapidly diagnose diseases. With the help of statistics, the approach avoids redundant fractionation procedures.

MALDI imaging mass spectrometry

Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry is known for its high sensitivity, ease of operation, and automated capabilities that can be used for the analysis of small organic or large biochemical compounds obtained from various sources.[1–6] MALDI-TOF employs a focused laser beam to desorb and ionize the sample, which is typically comprised of an analyte surrounded by a large amount of organic matrix. Due to the innate spatial resolution of the laser beam, the MALDI-TOF apparatus can be used to study the distribution of particular chemicals or biochemical compounds on a solid sample surface, such as a tissue. This method is known as imaging mass spectrometry (IMS), which is the ideal approach for determining the distribution of molecules of interest directly on the sample's surface without additional or extraneous handling.[7,8] Thus, IMS has emerged as a powerful tool for biomarker discovery and allows for the direct correlation between protein expression and histology.[9]

Based on its morphologically driven protocol, MALDI-based IMS allows the direct evaluation of different cells that are growing in a tissue. Moreover, unlike other imaging techniques such as fluorescence imaging or magnetic resonance imaging, MALDI-based IMS does not require prior knowledge of the sample before analysis. This technique can be used to determine the presence of hundreds of compounds in a single measurement without labeling. Fig. 1 shows the MALDI-based IMS experimental workflow. The protocol begins with cutting the frozen organ or tissue into sections ranging between 5—20 μm in thickness and thaw-mounting these sections onto stainless steel targets or Indium-Tin-Oxide (ITO) slides. A traditional MALDI matrix, such as 2,5-dihydroxybenzoic acid (2,5-DHB), -cyano-4-hydroxycinnamic acid (-CHC), or 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid), is then applied onto the tissue surface by an air brush or automatic matrix-preparation system. After a homogenous layer of fine matrix crystals is applied onto the tissue section, the target is transferred to the ion source of a MALDI-TOF mass spectrometer. A laser beam is then fired across the region of interest onto the tissue section using a set lateral resolution in order to generate ions for analysis by the TOF analyzer. The mass spectra are acquired and the imaging results are visualized by plotting the ion intensities as a function of the two-dimensional (2D) coordinates of the tissue using the imaging software.

To date, MALDI-based IMS has been used in a wide variety of applications. For example, drug metabolic processes were studied using IMS on whole body dissections of mice,[10] and the area growing within different types of cells on the tissue were well defined[11] and the accumulation of tainted chemical compounds in specific organs were proven using IMS.[12] The technique is a powerful tool for not only characterizing biomolecules in specific regions of the sample, but also their spatial distributions on the biological surface. In addition, it is clear that this approach can be integrated into clinical management for disease diagnosis and outcome prediction.

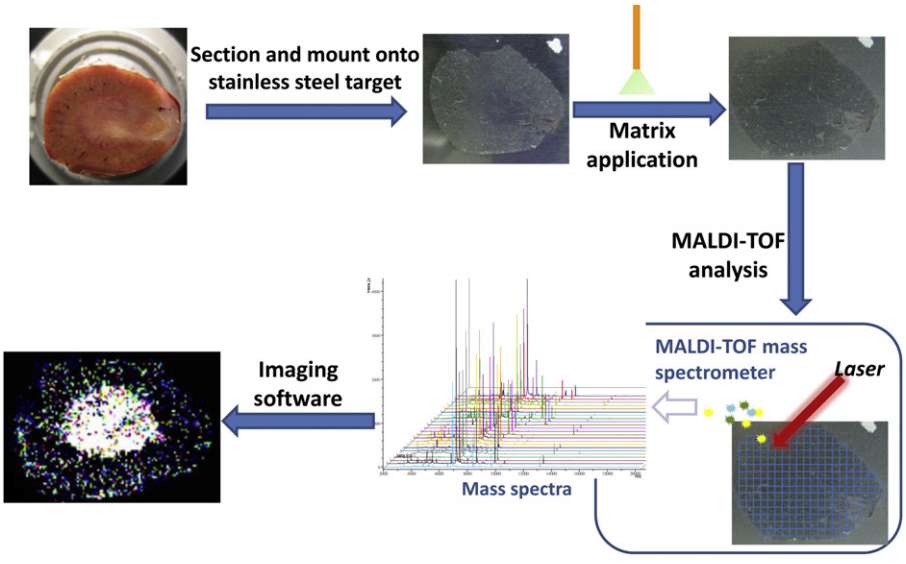

Figure 1 Experimental workflow of MALDI imaging. The MALDI imaging procedure is initiated by sectioning and mounting a tissue section on a stainless steel target, then depositing the matrix and irradiating a laser across the tissue surface. The mass spectra of the spots are set by a fixed lateral resolution, and the imaging results are visualized using imaging software.

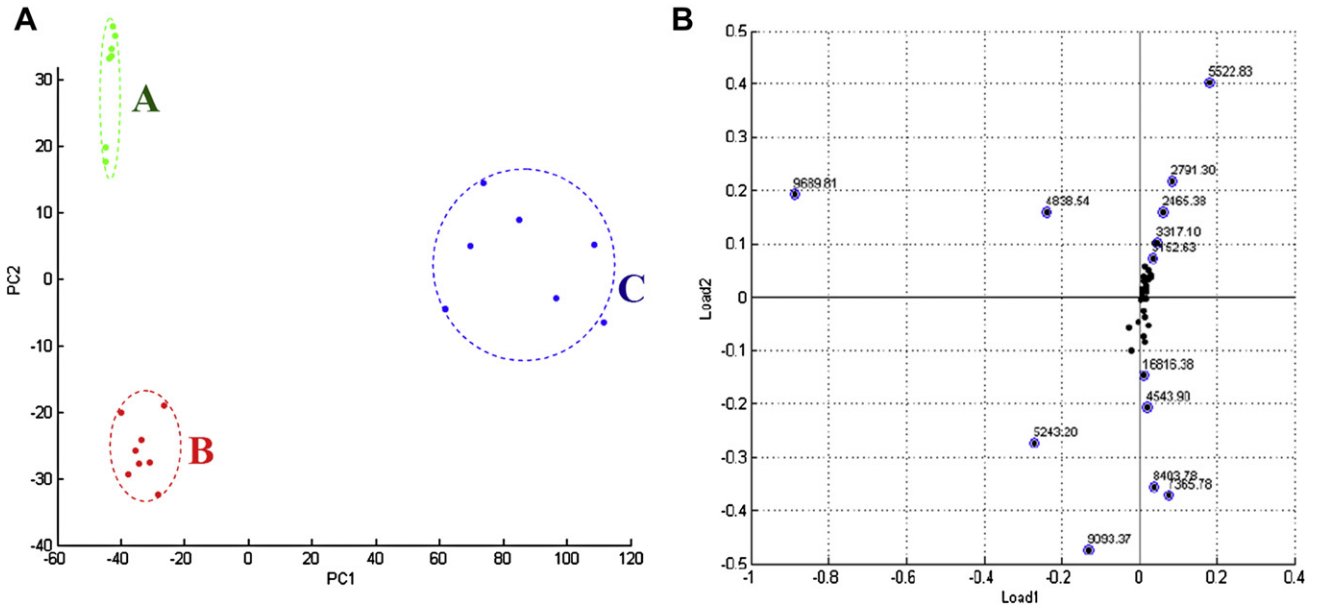

Figure 2 Representative plots of PCA results used to differentiate the extracts of three cell groups (A, B, and C) obtained from their MALDI mass spectra. (A) Score plot; (B) loading plot.

Principal component analysis for processing MALDI mass spectra data

In general, tissue specimens are highly complex systems that consist of lipids, peptides, proteins, and salts. The MALDI mass spectra obtained from these biological specimens provide information regarding the types of molecules in the sample. If MALDI data obtained from normal control specimens and patients can be well classified, grouped, and differentiated, it will be very helpful for the efficient diagnosis of disease. MALDI data are often used to differentiate lipids, expressed peptides, and proteins in comparative analysis, but efficiently comparing very complicated MALDI data has been a challenging issue for scientists.

Because MALDI mass spectra are a type of multivariate data, with each mass signal defining one molecular dimension, the process of evaluating mass spectra using multivariate statistical methods allows the straightforward differentiation of samples.Principle component analysis (PCA), a multivariate method designed to extract the variance within a data set, is one of the most widely used statistical methods used for differentiation. PCA reduces the dimensionality of the data set while retaining the information present in the original data set. By reducing the dimensionality of the data set to a 2D or 3D coordinate system, in which each sample (spectrum) is represented by a point, spectra with similar variation characteristics can be clustered together and the differences between sample groups can be readily visualized in the system. Because the first two principal components - PC1 and PC2 usually provide more than 80% of the total variance between the samples, it is sufficient to express the original multidimensional data matrix in a 2D plot, i.e., PC2 vs. PC1. PCA in combination with MS has been applied to the fields of metabolomics, proteomics, and MALDI-based IMSstudies. Fig. 2 shows the representative plots of the PCA results

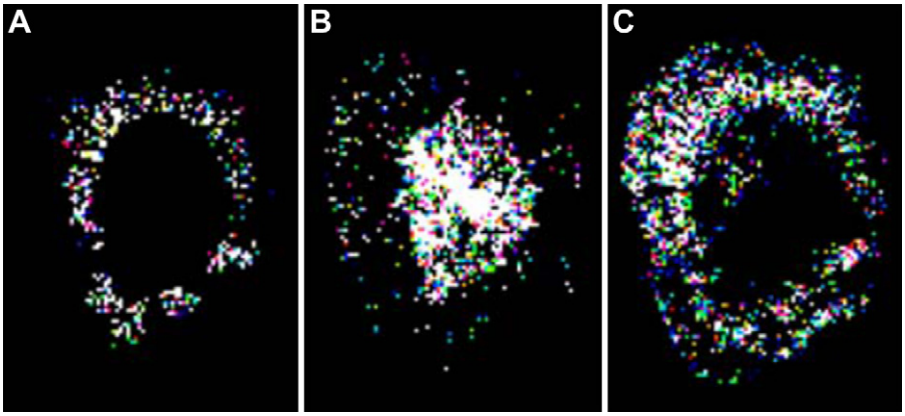

Figure 3 IMS images generated from signals at (A) m/z 741 (Sphingomyelin, SM 16:0 + K), (B) m/z 798 (Phosphatidylcholine, PC 34:1 + K), and (C) m/z 820 (Phosphatidylcholine, PC 36:4 + K) from renal samples obtained from a rat.

used for differentiating extracted MALDI mass spectra from three different cells. The PCA score plot indicates the similarities and differences between the ion patterns of the mass spectra of the three samples. The loading plots of the PCA analysis provide information regarding the contribution of each ion signal to the variance covered by each respective principle component.

The concept of using MALDI-TOF in combination with statistical processing to compare the lipid, peptide, or protein profiles of biological specimens, which carry abundant disease-related molecular information, can be used to rapidly diagnose diseases.[13] After the optimized experimental procedures are determined, samples can be analyzed using MALDI-TOF. PCA is then used to process all of the mass spectrometric data. This is different from the conventional approach where qualitative and quantitative data of specific biomarkers are used to diagnose certain diseases. Because the time required to complete MALDI-TOF and PCA analyses is usually short, these types of bioanalytical approaches for diagnosing diseases allows large-scale screening to become a reality.

Processing MALDI imaging data using PCA

MALDI-based IMS provides information on specific molecular distributions in a tissue. The information can then be correlated with clinical data on certain diseases. For example, it was found that the molecular patterns detected on the tissues of a developing tumor, invasive tumor, and normal epithelia are different. The molecular image based on specific signal markers can then be used to determine the areas with different states. Because MALDI-based IMS records a large number of mass spectra from a tissue section, finding the representative mass spectra and choosing the interesting peaks and biomarkers that display significant differences between tissues states have become important issues.Peak-by-peak interpretation of MALDI-based IMS data is a labor- and time-consuming process and inefficient. One of the efficient approaches is using PCA to process IMS data in order to reduce the complexity of the information of the multivariate data sets.[14,15] The grouped results in the score plot provide classification information regarding the tissue state. The loading plots of the PCA analysis can indicate the ion signals with significant differences between regions or tissues, and the imaging of these interesting peaks can be visualized by plotting the ion intensities across the tissue section. By using multivariate statistical methods, the chemical noise and the irrelevant features are eliminated and more concise information is obtained.

Previously, many studies used PCA to differentiate MALDI mass spectra recorded from different types of samples and process MALDI imaging data using PCA to diagnose disease. In our laboratory, several studies have been done using the same approach including: (1) the study of the distribution of melamine and lipids on rat renal tissues by IMS (Fig. 3)[12]; (2) rapid differentiation of the virulence of *Helicobacter pylori* using MALDI-TOF and PCA analyses; (3) combining MALDI-TOF and PCA for differentiating breast cancer cell lines with different Estrogen Receptor (ER) and Human Epidermal Growth Factor Receptor 2 (HER-2) statuses; and (4) applying MALDI-based IMS to define cancer regions in breast tissue. In conclusion, through the success of many studies, MALDI-TOF and MALDI-based IMS together with PCA have become techniques that are not only useful for academic research but also practical tools for clinical diagnosis.

Acknowledgments

This work was supported by grant from the Department of Health, Executive Yuan, R.O.C. (TAIWAN) (DOH100-TD-C-111-002).

References

1. Tanaka K, Hiroaki W, Yutaka I, et al. Protein and polymer analyses up to m/z 100000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom*. 1988;2:151—153.
2. Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10000 daltons. *Anal Chem*. 1988;60:2299—2301.
3. Fuchs B, Schiller J. Application of MALDI-TOF mass spectrometry in lipidomics. *Eur J Lipid Sci Technol*. 2009;111:83—98.
4. Su AK, Liu JT, Lin CH. Rapid drug-screening of clandestine tablets by MALDI-TOF mass spectrometry. *Talanta*. 2005;67:718—724.
5. Shiea J, Cho YT, Lin YH, et al. *Rapid Commun Mass Spectrom*. 2008;22:3754—3760.
6. Lin SY, Shih SH, Wu DC, et al. *Rapid Commun Mass Spectrom*. 2007;21:3311—3316.
7. Caprioli RM, Farmer TB, Gile J. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal Chem*. 1997;69:4751—4760.
8. Kristen DH, Stacey RO, Richard MC. *Semin Nephrol*. 2007;27:597—608.
9. Caprioli RM. Deciphering protein molecular signatures in cancer tissues to aid in diagnosis, prognosis, and therapy. *Cancer Res*. 2005;65:10642—10645.
10. Hsieh Y, Chen J, Korfmacher WA. Mapping pharmaceuticals in tissues using MALDI imaging mass spectrometry. *J Pharmacol Toxicol Methods*. 2007;55:193—200.
11. Walch A, Rauser S, Deininger SO, et al. MALDI imaging mass spectrometry for direct tissue analysis: a new frontier for molecular histology. *Histochem Cell Biol*. 2008;130:421—434.
12. Wang HY, Lo LH, Tyan YC. The study of the distribution of melamine in rat renal tissues by imaging mass spectrometry. *Anal Methods*. 2010;2:1974—1981.
13. Ragazzi E, Pucciarell S, Seraglia R, et al. Multivariate analysis approach to the plasma protein profile of patients with advanced colorectal cancer. *J Mass Spectrom*. 2006;41:1546—1553.
14. Djidja MC, Claude E, Snel MF, et al. Novel molecular tumour classification using MALDI-mass spectrometry imaging of tissue micro-array. *Anal Bioanal Chem*. 2010;397:587—601.
15. Yao I, Sugiura Y, Matsumoto M, et al. In situ proteomics with imaging mass spectrometry and principal component analysis in the Scrapper-knockout mouse brain. *Proteomics*. 2008;8:3692—3701.